

グラフカットを利用したステレオ画像からの 複数平面領域と平面パラメータの直接推定

藤原 将 展^{†1} 杉本 茂 樹^{†1} 奥 富 正 敏^{†1}

本論文では、ステレオ画像から直接、3次元シーン中に含まれる複数の平面領域および各平面のパラメータを同時に推定する手法を提案する。提案手法では、まず、3自由度の平面パラメータを離散化する。そして、各平面パラメータから算出される画像間の画素値の SAD と、平面領域の画像上での不連続性に関するペナルティ項を付与したコスト関数を定義し、そのコスト関数を、グラフカットの α 拡張アルゴリズムを用いて最小化することにより、複数の平面領域と、各平面のパラメータを同時に推定する。提案手法では、画像からの特徴点抽出や、事前のセグメンテーションといった別処理を必要とせず、入力された画像をそのまま利用して、複数の平面領域とそれらの平面パラメータを同時に推定できる。合成画像と実画像を用いた実験を通じて、提案手法の有効性を示す。

Direct Estimation of Multiple Planar Regions and Plane Parameters via Graph-Cuts Using Stereo Images

MASANOBU FUJIWARA,^{†1} SHIGEKI SUGIMOTO^{†1}
and MASATOSHI OKUTOMI^{†1}

In this paper, we propose a method for simultaneously estimating multiple planar regions and the plane parameter of each plane directly from stereo images. In the proposed method, we first discretize plane parameter. Then, by using the Graph-Cuts α -expansion algorithm, we minimize a cost function composed by (1) a data term representing the pixel value dissimilarity between the two images one of which is warped by the homography computed from each plane parameter, and (2) a smoothness constraint term representing the penalty on planar region discontinuity. The proposed method can obtain multiple planar regions and their plane parameters along with non-planar regions by using an upper bound for the pixel value dissimilarity. The validity of the proposed method is demonstrated through experiments using synthetic and real images.

1. はじめに

人工的な環境は多数の平面によって構成されている。よって、画像を用いてこれらの平面を抽出し、その位置や姿勢を推定することができれば、屋内環境の3次元モデリングや、移動ロボットのための周辺環境認識など、様々なアプリケーションに適用できる。しかし、画像中の平面領域と、その平面の3次元空間中の位置・姿勢を示すパラメータとを同時に推定する問題は、基本的に Chicken-and-Egg 問題であり、既存手法では、一方を仮定して他方を得るというアプローチが一般的であった。

各画素について平面領域に属するか否かを判断する手法としては、単眼動画画像やステレオ画像を利用して平面射影変換を推定した後に、一方の画像を変換することにより2枚の画像を重ね合わせ、画素値の差分の小ささに基づいて平面領域を抽出する手法が提案されている^{1),4),8),10)-12)}。例えば、ブロック単位で差分値を比較する手法^{4),10)}や、差分画像を二値画像化して拡大縮小処理を行う手法^{11),12)}などがある。ただし、このような処理方法では、平面と非平面との境界を精度よく推定することは困難である。

一方、画像セグメンテーション手法として、隣接画素の類似性と境界の平滑性をコスト関数で表現し、それをグラフカットによって最小化する方法が知られている^{6),9),15)}。そして、1枚の画像における隣接画素間の画素値の類似性の代わりに、動画画像によるトラッキング結果を利用して、画像を重ね合わせた際の画素値の類似性を利用することにより、グラフカットを利用して動画画像から動物体をセグメンテーションする方法が提案されている^{2),14)}。ただし、このような方法では、別処理として、動画画像のトラッキング結果からの画像の変換パラメータの推定が必要な上、セグメンテーション結果はトラッキング精度に大きく依存する。

これに対し、本論文では、ステレオ画像から直接、3次元シーン中に含まれる複数の平面領域および各平面のパラメータを同時に推定する手法を提案する。提案手法では、図1に示すように、まず、3自由度の平面パラメータを3次元空間内で離散化する。そして、各平面パラメータから算出される画像間の画素値の SAD (Sum of Absolute Difference) と、平面領域の画像上での不連続性に関するペナルティ項を付与したコスト関数を定義し、そのコスト関数を、グラフカットの α 拡張アルゴリズムを用いて最小化することにより、複数の平面領域と、各平面のパラメータを同時に推定する。

^{†1} 東京工業大学大学院理工学研究科機械制御システム専攻
Graduate School of Science and Engineering Tokyo Institute of Technology

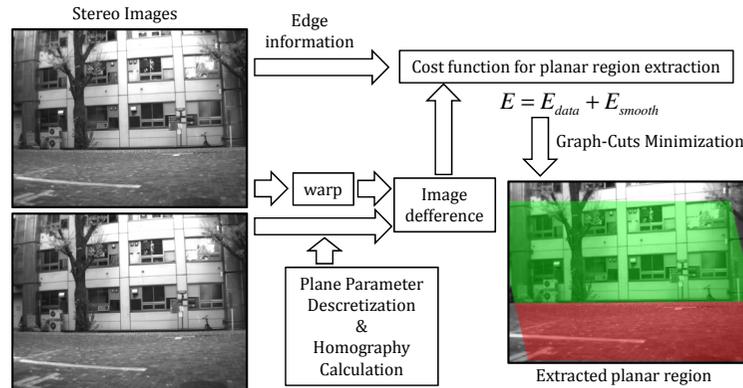


図 1 提案手法の流れ

Fig.1 Flow of the overall process

提案手法では、画像からの特徴点抽出や、事前のセグメンテーションといった別処理を必要とせず、入力された画像をそのまま利用して、複数の平面領域とそれらの平面パラメータを同時に推定できる。また、画像間の差分絶対値の最大値を定義することで、どの平面にも属さない画像上の領域も同時に抽出することができ、これにより、地平面のみを推定したい場合には、鉛直方向の法線周りのみパラメータを離散化すればよいなど、目的に適した応用が可能である。

本論文の構成は以下の通りである。まず、2節において、提案手法のアプローチと、本文中で用いる平面パラメータおよび平面パラメータの離散化方法について述べ、3節において、推定に用いるコスト関数の設計方法と、グラフカットおよび α 拡張アルゴリズムを用いたコスト関数の最適化方法について述べる。そして、4節において、合成画像と実画像を用いた実験を通じて提案手法の有効性を示し、最後に5節において本論文のまとめについて述べる。

2. 提案手法の概要および平面パラメータの離散化

2.1 提案手法のアプローチ

提案手法は、シーン中に含まれる複数の平面領域および各平面パラメータを推定する。しかし、特徴点抽出や3次元計測などの結果から事前に平面パラメータの候補を推定する既存手法^{2),7),14)}とは異なる。本手法においても、最終的にグラフカットを用いたコスト関数の

最適化により、各画素に平面の候補ラベルをセグメンテーションするが、用いる平面の候補については、3自由度の平面パラメータ空間を離散化することで生成する。生成された多数の平面パラメータそれぞれを平面の候補ラベルとし、各画素がどのラベルに属するかを決定することにより、平面領域と各平面パラメータを同時に推定する。平面パラメータの離散化方法については後述する。

平面パラメータを用いて、ステレオ画像のうち一方の画像を平面射影変換し、基準画像との間で画素値の差分を計算すると、平面パラメータが正しい場合には、その差分値は小さくなる。よって多数の平面の候補を各画素に対して割り当てる際には、この画素値差分の小ささを基準にすればよいと考えられる。しかし、1画素ごとに画素値差分を計算すると、平面の候補の中でその値を小さくするような平面パラメータは多数存在する。そこで本手法では、隣接画素間で同一の平面に属するという平滑化拘束を加える。平滑化拘束を与えることで、複数の画素をまとめた領域全体において画素値差分の小ささを評価していることになり、曖昧性を減少させることで平面の候補を限定することができる。これらの考え方に基づくコスト関数について、3.2節において詳しく述べる。また、このようなコスト関数は、グラフカットの α 拡張アルゴリズム^{3),9),15)}を利用することで最適化することが可能である。

2.2 平面パラメータ

平面パラメータとは、3次元空間中の平面の方程式を表す3自由度のパラメータであり、次式のように3次元のベクトル \mathbf{q} で表す。

$$\mathbf{q} = (q_x, q_y, q_z)^T \quad (1)$$

このとき、平面の方程式は次式のようにになる。

$$q_x x + q_y y + q_z z = 1 \quad (2)$$

また本論文では、この平面パラメータを次式のようなパラメタライゼーションにより表現する。

$$\mathbf{q} = \frac{\mathbf{n}^T}{d} = \left(\frac{\cos \psi \cos \theta}{d}, \frac{\cos \psi \sin \theta}{d}, \frac{\sin \psi}{d} \right)^T \quad (3)$$

ただし、 \mathbf{n} は平面の法線ベクトル、 d は光学中心から平面におろした垂線の長さ、つまり光学中心から平面までの距離を表し、 ψ および θ は法線ベクトルの三角関数表現における角度を表す2つのパラメータである。図2に、法線ベクトルを示す。

2.3 平面パラメータの離散化方法

本手法では、3自由度の平面パラメータを離散化し、図3のように多数の平面を生成し、生成されたそれぞれの平面を、入力画像中の各画素が属する平面の候補のラベルとして保

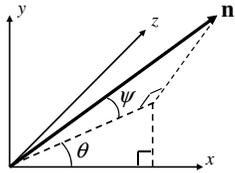


図2 法線ベクトル
Fig.2 Normal vector

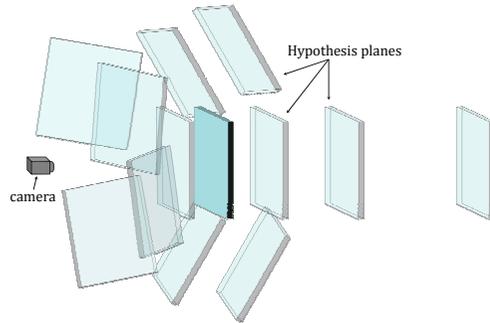


図3 平面の候補群
Fig.3 Hypothesis planes

持する。2.2節で示した通り、3次元空間中で平面パラメータは、3自由度のパラメータ ψ , θ , d を用いて式(3)のように表現される。これを次式のように表現しなおす。

$$\mathbf{q} = \mathbf{n}^T \cdot \bar{d} = (\bar{d} \cdot \cos \psi \cos \theta, \bar{d} \cdot \cos \psi \sin \theta, \bar{d} \cdot \sin \psi)^T \quad (4)$$

ただし、 \bar{d} は平面までの距離の逆数 ($\bar{d} = 1/d$) を表し、3つのパラメータ ψ , θ , \bar{d} で平面パラメータを表現する。そして、それぞれを ψ_d , θ_d , \bar{d}_d という固定間隔で離散化する。平面までの距離を、逆数として表現し離散化するのは、平行ステレオカメラに正対するフロント平行な平面を仮定した際に、その平面に対するステレオ画像間の視差が平面までの距離に反比例するためである。

2.4 対象シーンに適した効率的な離散化

本手法では、多数の平面の候補ラベルを用いて推定を行うが、3自由度の平面パラメータを離散化した場合、ラベル数のオーダーは、離散化間隔に対して $O(n^3)$ となる。処理の安定化や高速化のためには、このラベル数を抑えることが重要である。提案手法では、対象シーンに関する事前知識を用いて離散化の範囲を限定してラベル数を抑える。

道路シーンや建造物、屋内シーンなどにおいては、鉛直上向きに近い法線ベクトルを持つた地平面が存在することが仮定できる。また、車載カメラや移動ロボットの周辺環境認識などのアプリケーションを想定した場合、地平面とカメラの位置関係はおおよそ既知であると考えられる。よって、地平面のみを抽出したい場合、平面パラメータの各パラメータのうち、法線ベクトルを表す ψ, θ に関しては、おおよその鉛直方向周辺のみを離散化し、さらに平面までの距離の逆数 \bar{d} に関する離散化を行い、平面の候補ラベルを生成すればよい。

このような離散化を行うことで、ラベル数が限定される。そして、カメラから平面までの距離に関しても値がおおよそ既知であるアプリケーションを想定した際には、このラベル数を更に抑えることができる。

また、上記のようなシーンでは、地平面以外の平面は、地平面に垂直な平面で構成されることが多いと考えられる。よって、地平面以外の平面も抽出したい場合には、まず地平面の推定を行い、その後、推定された地平面に加えて、地平面に垂直な法線ベクトルをもつ平面パラメータの離散化を行い平面の候補ラベルを生成し、再び推定を行えばよい。

このように、推定対象とする平面を、地平面および地平面に垂直な平面に限定することで、平面の候補ラベルの数を抑えることが可能である。4節の実験では、鉛直方向がおおよそ既知であることを仮定して地平面を抽出した後、それに加えて地平面に垂直な法線ベクトルを持つ平面パラメータの離散化を行い、地平面および地平面に垂直な平面を抽出した結果を示す。

3. コスト関数の設計および最適化

3.1 ステレオ画像間の平面射影変換行列

本手法では、最終的にコスト関数を設計し、それをグラフカットを利用して最適化することにより、2節で述べた平面パラメータの離散化で生成された多数の平面の候補ラベルを各画素に割り当て、平面領域と平面パラメータの推定を行う。ここでは、コスト関数を設計する際に必要な、ステレオ画像間の平面射影変換行列について説明する。

基準画像座標を $\mathbf{u} = (u, v)^T$, 参照画像座標を $\mathbf{u}' = (u', v')^T$ とする。いま、ある平面を2枚の画像上に射影したときの2枚の画像の座標関係は、同次座標表現 $\bar{\mathbf{u}}, \bar{\mathbf{u}'}$ を用いて、
$$\bar{\mathbf{u}} \approx \mathbf{H} \bar{\mathbf{u}'} \quad (5)$$
 と書くことができる。ただし、 \mathbf{H} は 3×3 の平面射影変換行列を表している。このとき、平面射影変換行列 \mathbf{H} と平面パラメータベクトル \mathbf{q} との関係は、次式のように表される⁵⁾

$$\mathbf{H} = \mathbf{A}_1 (\mathbf{R} + \mathbf{t} \mathbf{q}^T) \mathbf{A}_0^{-1} \quad (6)$$

ただし、 $\mathbf{A}_0, \mathbf{A}_1$ は、それぞれ基準カメラと参照カメラの内部パラメータ行列を表し、 \mathbf{R}, \mathbf{t} はステレオカメラ間の回転と並進の関係を表すパラメータである。

この平面射影変換行列を、離散化されたそれぞれの平面ごとに計算し、その結果を利用してコスト関数を設計する。

3.2 平面領域抽出のためのコスト関数

ある画素 \mathbf{u} が、どの平面領域に属するのかを表す属性 (ラベル) を $S_{\mathbf{u}}$ とし、その値を以下のように定める。なお、ラベル $S_{\mathbf{u}} = 0$ は、どの平面にも属さない非平面ラベルを表す。

$$S_{\mathbf{u}} = \begin{cases} l & : \text{画素 } \mathbf{u} \text{ が平面領域に属するとき } (l = 1, 2, \dots) \\ 0 & : \text{画素 } \mathbf{u} \text{ が平面領域に属さないとき} \end{cases} \quad (7)$$

そして、全画素に関するラベルの集合を $\mathbf{S} = \{S_{\mathbf{u}}\}$ (画素数の次元のもつベクトル) とし、 \mathbf{S} によって定まるコスト $E_{region}(\mathbf{S})$ を次式で定義する。

$$E_{region}(\mathbf{S}) = \sum_{\mathbf{u} \in ROI_r} E_{data}(S_{\mathbf{u}}) + \kappa \sum_{(\mathbf{u}, \mathbf{v}) \in N} E_{smooth}(S_{\mathbf{u}}, S_{\mathbf{v}}) \quad (8)$$

式 (8) の第一項 E_{data} はデータ項であり、画素 \mathbf{u} の属性 $S_{\mathbf{u}}$ のみによって定まるコストを表す。この項は、画素 \mathbf{u} が $S_{\mathbf{u}}$ の属性を持つ確率が高いほどコストが小さくなるように設計する。なお、 ROI_r は画像をセグメンテーションする領域を表す。データ項 E_{data} は具体的には以下のように定める。

$$E_{data} = \begin{cases} |I_0[\mathbf{u}] - I_1[\mathbf{w}(\mathbf{u}; \mathbf{H}_{S_{\mathbf{u}}})]| & : \text{if } S_{\mathbf{u}} \neq 0 \\ E_{NP} & : \text{if } S_{\mathbf{u}} = 0 \end{cases} \quad (9)$$

ただし、 I_0, I_1 はそれぞれ基準画像と他方の画像の画素値を表し、 \mathbf{u} は基準画像上の座標を表す。また、 $\mathbf{H}_{S_{\mathbf{u}}}$ は、平面ラベル $S_{\mathbf{u}}$ に対応した、8 自由度の平面射影変換行列である。 $\mathbf{w}(\mathbf{u}; \mathbf{H}_{S_{\mathbf{u}}})$ は $\mathbf{H}_{S_{\mathbf{u}}}$ に依存した平面射影変換関数であり、座標 \mathbf{u} を変換した座標値を表す。 E_{NP} は非平面領域に対する定数コストである。

このコストは、閾値処理に基づいた従来手法^{1),4),8),10)-12)} の考え方に基いており、式 (9) の E_{NP} がその閾値の役割を果たしている。すなわち、画素 \mathbf{u} における差分値 $|I_0[\mathbf{u}] - I_1[\mathbf{w}(\mathbf{u}; \mathbf{H}_{S_{\mathbf{u}}})]|$ が E_{NP} よりも小さければ、 $S_{\mathbf{u}} \neq 0$ となり、そうでなければ、 $S_{\mathbf{u}} = 0$ となるように属性が定められる。また、もし画素 \mathbf{u} に割り当てられたラベル $S_{\mathbf{u}}$ が正しい平面ラベルであれば、この値は小さくなるのが期待できる。

しかし、このデータ項 $E_{data}(S_{\mathbf{u}})$ の値は、キャリブレーション誤差や平面パラメータの離散化間隔の粗さに応じた誤差、また対象の非平面性等の影響により、参照画像の変換を行い基準画像との画素値の差分を取った際に残差が生じ、特にこの残差の値は、画像上の輝度のエッジ部分において顕著に表れる。よって、実際には3次元空間中で同一平面上の領域であっても、画像上のセグメンテーション結果は輝度のエッジ部分において、ラベルの変化が

起こりやすくなってしまい、所望の推定結果が得られない場合がある。そこで本手法では、データ項 $E_{data}(S_{\mathbf{u}})$ の計算の際に、基準画像上の座標 \mathbf{u} を平面射影変換行列 $\mathbf{H}_{S_{\mathbf{u}}}$ で変換した座標値 $\mathbf{w}(\mathbf{u}; \mathbf{H}_{S_{\mathbf{u}}})$ における画素値との差分値だけでなく、近傍の複数の画素値との差分値も計算し、それらの差分値のうち最小の値をデータ項 $E_{data}(S_{\mathbf{u}})$ の値とすることで、上記の影響を抑える。よって、本手法では、式 (9) を書き換えた次式を、データ項 $E_{data}(S_{\mathbf{u}})$ として用いる。

$$E_{data}(S_{\mathbf{u}}) = \begin{cases} \min_{\mathbf{u}_{neighbor} \in N_{data}} |I_0[\mathbf{u}] - I_1[\mathbf{u}_{neighbor}]| & : \text{if } S_{\mathbf{u}} \neq 0 \\ E_{NP} & : \text{if } S_{\mathbf{u}} = 0 \end{cases} \quad (10)$$

ただし、 N_{data} は近傍画素の集合を表しており、小数位置で表される変換後の座標値 $\mathbf{w}(\mathbf{u}; \mathbf{H}_{S_{\mathbf{u}}})$ および $\mathbf{w}(\mathbf{u}; \mathbf{H}_{S_{\mathbf{u}}})$ を取り囲む整数位置の座標値を用いている。 $\mathbf{u}_{neighbor}$ は近傍画素を表す変数である。この近傍画素の集合 N_{data} は、そのサイズ λ が調整パラメータとなっており、図 3.3 に示すように、 $\lambda = 1$ の場合は周囲を取り囲む 4 画素、 $\lambda = 2$ の場合は周囲を取り囲む 16 画素がその範囲となる。この λ の値は、上述のキャリブレーション誤差や平面パラメータの離散化間隔の粗さに応じた誤差、対象の非平面性等の影響を、どの程度許容するかを調整するパラメータとなっている。

式 (8) の第二項である平滑化項 E_{smooth} は、平面領域での近傍画素間の連続性と、エッジによる境界の不連続性を考慮したペナルティとなっている。なお、 N は近傍画素の集合 (本論文では 8 近傍) を表し、 κ は平滑化項の重みを定める係数である。連続性は、ある画素 \mathbf{u} が属性 $S_{\mathbf{u}}$ を持つならば、その近傍の画素も同じ属性を持つ可能性が高いことを示し、エッジによる不連続性は、画素値が異なる場合には異なる属性を持つ可能性が高いことを示す。提案手法では、これらの性質を次式のようなコストによって表現する。

$$E_{smooth} = \begin{cases} 0 & : \text{if } S_{\mathbf{u}} = S_{\mathbf{v}} \\ \frac{1}{D(\mathbf{u}, \mathbf{v}) \cdot (|I_0[\mathbf{u}] - I_0[\mathbf{v}]| + C)} & : \text{if } S_{\mathbf{u}} \neq S_{\mathbf{v}} \end{cases} \quad (11)$$

ただし、 $D(\mathbf{u}, \mathbf{v})$ は、画素 \mathbf{u} と画素 \mathbf{v} との間の距離を表し、 C は分母が零になることを防ぐための定数である。すなわち、提案手法では、隣接画素の属性が異なるときに距離に応じたコストを与え、一方エッジがある場合には与えるコストを小さくしている。

3.3 コスト関数の最適化

提案手法のコスト関数 (8) はグラフカットによって最適化することができる^{6),9),15)}。ラ

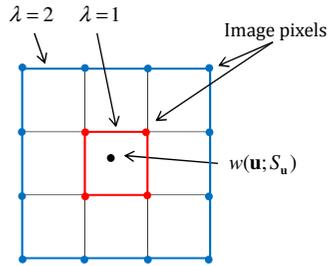


図4 周辺画素
Fig. 4 Surrounding pixels

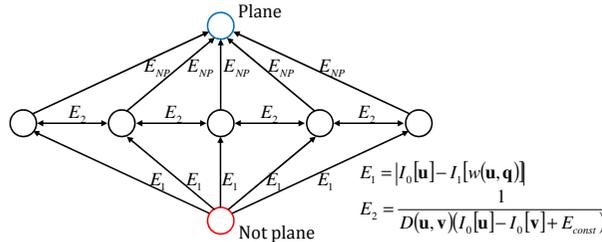


図5 グラフの構造
Fig. 5 Graph structure

ベルの数が2つの場合、つまり平面領域と非平面領域への分割を考えると、式(8)により、図5に示す構造のグラフを構成することができ、そのmax-flowを求めることで、各画素の最適なラベル付けを得ることができる。

図5の黒いノードはそれぞれ基準画像上の画素を表している。また、青いノードは平面領域への属性を表している。赤いノードと黒いノードを結ぶ実線にそれぞれ平面に属する場合のコストを設定し、青いノードと黒いノードを結ぶ実線にはそれぞれ閾値 E_{NP} をコストとして設定する。そして、黒いノード間を結ぶ実線には、平滑化項の値を設定する。このようにして作成されたグラフをグラフカットアルゴリズムによって切断することにより、式(8)のコストを最小化する属性を得ることができ、平面領域が抽出される。

しかし、グラフカットではラベル数が2つの場合には、上記のようにmax-flowを求めることでコスト関数の最小化を行うことができるが、ラベル数が3つ以上の場合には α 拡張³⁾などの、多値ラベルに対する近似最小化アルゴリズムを適用する必要がある。 α 拡張では、2ラベルのグラフカット問題を繰り返し解くことで、近似的に多値ラベルのコスト関数の最小化を行う。

α 拡張では、まず任意のラベル付け S から始め、繰り返し移動をする。一回の移動、つまりラベル付け S の変更は、1つのラベル α を固定しておいて、各画素について現在のラベル付けをそのままにするか、 α に変更するかを決定することにより行われる。このように、ラベルを α に変更するかどうかの選択はラベルが2つのグラフカット問題として定義され、このグラフカット問題を全てのラベル α について繰り返し解くことにより、 $E_{region}(S)$ を近似的に最小化する。

α 拡張のフローを以下に示す。なお、ラベル全体の集合を L と表す。

- (1) Start with an arbitrary labeling S
- (2) Set success := 0
- (3) For each label $\alpha \in L$
 - (3-1) Find $\hat{S} = \arg \min E_{region}(\hat{S})$ among \hat{S} within one α -expansion of S
 - (3-2) If $E_{region}(\hat{S}) < E_{region}(S)$ set $S := \hat{S}$ and success := 1
- (4) If success = 1 goto 2
- (5) Return S

このように、 α 拡張を用いることにより、複数ラベルのコスト関数を近似的に最小化することができる。しかし、 α 拡張の計算量は、ラベル数に応じて増加するため、本手法のように多数のラベルを扱う場合、計算時間を短縮する工夫が必要である。そこで本手法では、さらに階層的なアルゴリズムを利用する。3.4節で階層的アルゴリズムについて説明する。

3.4 階層的アルゴリズム

3.3節で述べた通り、本手法では、計算時間を短縮し処理を高速化するために、階層的なアルゴリズムを用いる。

まず、画像をダウンサンプリングし、ガウシアンピラミッドを作成し、最も解像度の低い階層において提案手法を適用する。次に、推定結果において、閾値以上の画素数に対して割り当てられている平面ラベルのみを残し、それ以外のラベルを削除する。次いで、残った全ての平面ラベルに対応する平面パラメータの周辺のみを、前回の離散化時よりも細かく離散化し直し、新たな平面ラベルとして保持する。そして、新たに作成した平面ラベルを用いて、1階層高解像度の画像に対して提案手法を適用する。このように、推定およびラベルの再構成を繰り返し、最終的に原画像の各画素に対して推定を行う。

グラフカットの α 拡張アルゴリズムにおける計算時間は、画素数に応じて増加するため、解像度の低い層で大きなラベル数に対する推定を行うことにより、計算時間が大幅に短縮される。そして、解像度の低い層での推定時には、粗い離散化間隔で構成された平面ラベル群を用いることにより、平面ラベル数を抑えることができ、これによっても計算時間が短縮される。また、階層が変わる際には画像が高解像度になるため、離散化間隔を細分化し、平面ラベルを再構成することによって、平面パラメータの推定精度を保つ。

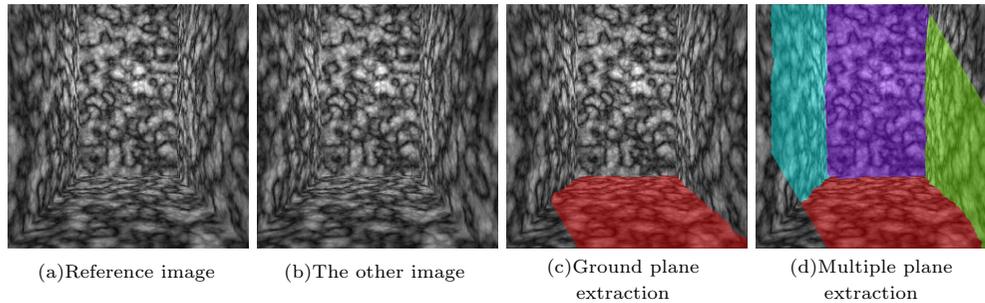


図 6 合成画像に対する平面抽出結果
Fig. 6 Extraction results for synthetic images

平面	推定値	真値
地平面	$(-0.0056, 0.5124, -0.0056)^T$	$(0, 0.5, 0)^T$
左	$(-0.4749, -0.0051, 0.0104)^T$	$(-0.5, 0, 0)^T$
右	$(0.4748, 0.0053, 0.0103)^T$	$(0.5, 0, 0)^T$
正面	$(-0.0000, 0.0011, 0.0999)^T$	$(0, 0, 0.1)^T$

表 1 平面パラメータ推定結果
Table 1 Plane parameter estimation

	階層なし	階層あり
地平面	1.5×10^5 [s]	52.42[s]
複数平面	3.5×10^5 [s]	205.89[s]
合計	5.0×10^5 [s]	258.31[s]

表 2 階層的処理による高速化
Table 2 Computational time reduction by pyramidal processing

4. 実験結果

4.1 合成画像実験

提案手法を合成画像実験に適用した。使用したステレオ画像は、図 6(a), (b) に示す 512×512 pixel のグレイスケール画像であり、地平面の他、左右および正面の計 4 平面で構成されるシーンである。地平面抽出の際の平面パラメータの離散化範囲は、それぞれ $-15^\circ \leq \psi \leq 15^\circ$, $75^\circ \leq \theta \leq 105^\circ$, $0.1m^{-1} \leq \bar{d} \leq 3m^{-1}$ であり、離散化間隔は、 $\psi_d = 5^\circ$, $\theta_d = 5^\circ$, $\bar{d}_d = 0.1m^{-1}$ を用いる。ただし、階層的な処理において、階層が変わる度に離散化間隔の細分化を行う際、離散化間隔を $1/2$ としており、本実験では階層数を元画像を含めて 3 に設定しているため、最終的に離散化の間隔は $1/4$ となる。複数平面の抽出時には、2.4 節で述べた通り、推定された地平面に垂直な平面パラメータの離散化を行う。なお、平面パラメータの離散化および、階層的な処理の際の細分化において、真値の平面パラメータを持つラベルは生成されないようにしている。

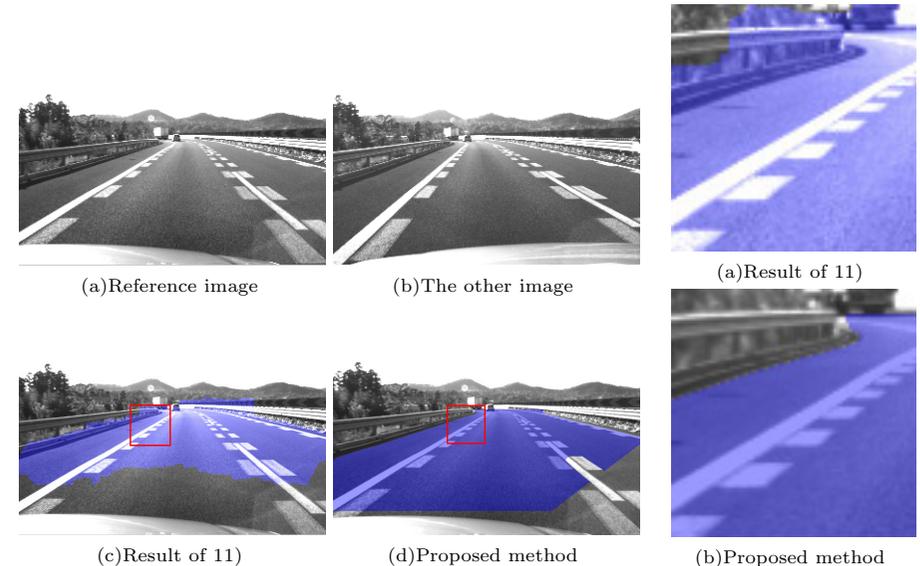


図 7 従来手法 11) と提案手法による地平面抽出結果
Fig. 7 Result of 11) and proposed method

図 8 従来手法 11) との比較 (拡大)
Fig. 8 Comparison with result of 11) (enlarged)

地平面の抽出結果を図 6(c) に、地平面および地平面に垂直な平面の抽出結果を同図 (d) に示す。さらに、表 1 に、推定された平面パラメータを真値とともに示す。画像上の平面領域および各平面パラメータの値が良好に推定されていることがわかる。また、表 2 に、3.4 節で述べた階層的な処理を用いた場合と用いなかった場合の計算時間の違いについて示す。なお、階層的な処理を用いなかった場合の計算時間については、少ないラベル数で行った実験での計算時間より見積もったおおよその値である。階層的な処理を用いることにより、大幅に計算時間が短縮化されることがわかる。

4.2 実画像実験

次に、実画像実験に対し提案手法を適用した結果を示す。まず、閾値処理による二値画像化と拡大縮小処理を行う従来手法¹¹⁾と、提案手法における地平面の抽出結果を比較した。なお、利用したステレオ画像は、 640×480 pixel のグレイスケール画像であり、離散化の範囲および間隔は、合成画像実験と同様の値を用いている。

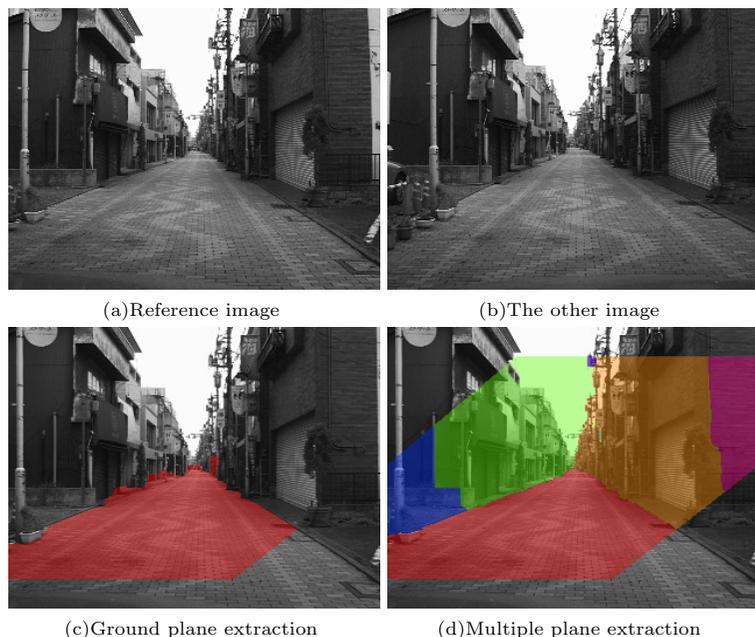


図 9 複数平面の抽出結果
Fig. 9 Result of multiple plane extraction

図 7 に、使用したステレオ画像 (同図 (a) と (b)) と、2 つの手法による平面領域抽出結果 (同図 (c) と (d)) を示す。また図 7(c), (d) に示した矩形領域を拡大したものを、図 8 に示す。従来手法による結果では、平面の境界付近での平面外領域への乗り上げや、道路上の白線部分での穴あきが見られる。それに対し、提案手法では、乗り上げや穴あきは比較的少なく、良好な抽出結果となっていることがわかる。

次に、地平面および地平面に垂直な平面の抽出結果を示す。図 9 に、使用したステレオ画像 (同図 (a) と (b)) と、地平面の抽出結果 (同図 (c))、地平面および平面に垂直な平面の抽出結果 (同図 (d)) を示す。また同様に、他のシーンに対しての推定結果を図 10 に示す。入力画像に対して、複数の平面領域が良好に抽出されていることがわかる。

また、図 9 および図 10 に示した結果では、複数平面の抽出結果において、地平面領域の抽出結果が改善されている。まず、乗り上げに関しては、地平面外領域への乗り上げの見ら

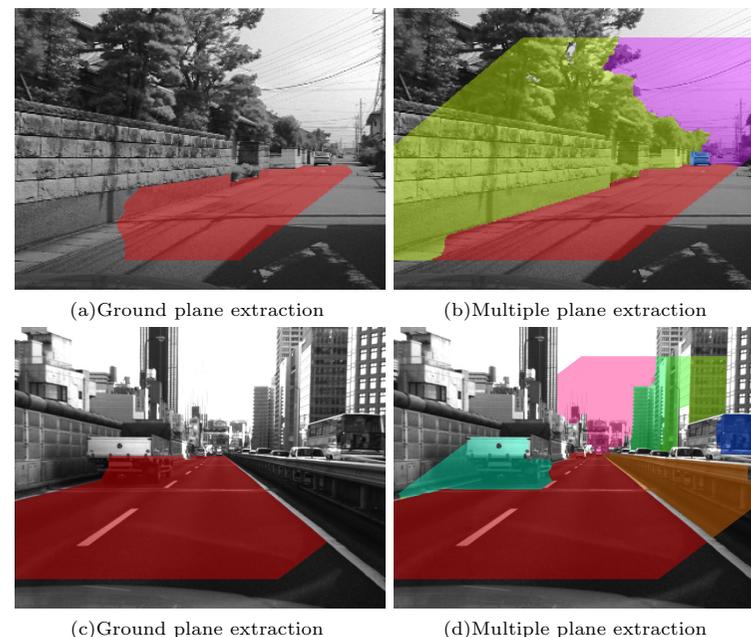


図 10 その他のシーンにおける複数平面の抽出結果
Fig. 10 Result of multiple plane extraction for other scenes

れた領域において、よりコストが小さい値となる別の適切な平面パラメータが割り当てられることにより改善される。そして、穴あきに関しては、乗り上げが起きにくくなることにより、平滑化項の重みをより大きな値に設定でき、その結果、推定結果が改善されると考えられる。

5. おわりに

本論文では、グラフカットを用いて、複数の平面領域および各平面パラメータを同時に、ステレオ画像から直接推定する手法を提案した。本手法では、3 自由度のパラメータで表現される平面パラメータの各パラメータを離散化し、各画素が属する平面の候補ラベルとして、多数の平面を生成した。そして、各画素へのラベル付け問題を、コスト関数の最適化問題として表すことで、平面パラメータの候補を推定するための別処理を必要とする既存手法

とは異なり，平面領域および各平面パラメータを同時に，ステレオ画像から直接推定した．実験結果では，合成画像および実画像に対し，提案手法を適用した結果を示し，画像中の複数の平面領域を良好に抽出できることを示した．

コスト関数の最適化に用いているグラフカットは，比較的時間がかかる処理になっている．提案手法では，このグラフカットによる最適化をラベル数に応じた回数行う必要があるため，処理に時間を要する．GPUなどのハードウェアを利用するなどの方法¹³⁾で，処理の高速化について検討する予定である．また，抽出結果に対して平面姿勢推定¹⁶⁾を適用して結果を改善することや，時系列ステレオ動画をを用いた高速カメラモーション推定¹⁷⁾を利用して，大規模なシーンの3次元形状復元なども検討する予定である．

参 考 文 献

- 1) S.Baker, R.Szeliski, and P.Anandan. A layered approach to stereo reconstruction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 434–441, 1998.
- 2) P. Bhat, K.C. Zheng, N. Snavely, A. Agarwala, M. Agrawala, M.F. Cohen, and B. Curless. Piecewise image registration in the presence of multiple large motions. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 2491–2497, 2006.
- 3) Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.23, No.11, pp. 1222–1239, 2001.
- 4) F. Dornaika and A.D. Sappa. Real Time Stereo Image Registration for Planar Structure and 3D Sensor Pose Estimation. *InTech Education and Publishing*, pp. 299–316, 2008.
- 5) O. Faugeras and F. Lustman. Motion and structure from motion in piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.2, No.3, pp. 485–503, 1988.
- 6) D. Freedman and P. Drineas. Energy minimization via graph cuts: Settling what is possible. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 939–946, 2005.
- 7) D. Gallup, J.M. Frahm, and M. Pollefeys. Piecewise planar and non-planar stereo for urban scene reconstruction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1418–1425, 2010.
- 8) M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision*, Vol.12, No.1, pp. 5–16, 1994.
- 9) V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.26, No.2, pp. 147–159, 2004.
- 10) H. Nakai, N. Takeda, H. Hattori, Y. Okamoto, and K. Onoguchi. A practical stereo scheme for obstacle detection in automotive use. In *International Conference on Pattern Recognition*, Vol.3, pp. 346–350, 2004.
- 11) M. Okutomi, K. Nakano, J. Maruyama, and T. Hara. Robust estimation of planar regions for visual navigation using sequential stereo images. In *IEEE International Conference on Robotics and Automation*, Vol.4, pp. 3321–3327, 2002.
- 12) A. Seki and M. Okutomi. Robust obstacle detection in general road environment based on road extraction and pose estimation. In *IEEE Intelligent Vehicles Symposium*, pp. 437–444, 2006.
- 13) V. Vineet and P. J. Narayanan. CUDA cuts: Fast graph cuts on the GPU. *CVPR Workshop on Visual Computer Vision on GPUs*, 2008.
- 14) J. Wills, S. Agarwal, and S. Belongie. What Went Where. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.1, pp. 37–454, 2003.
- 15) 石川博. グラフカット (チュートリアル). 情報処理学会研究報告 (コンピュータビジョンとイメージメディア 2007-CVIM-158), No.31, pp. 193–204, 2007.
- 16) 杉本茂樹, 奥富正敏. ステレオ画像を利用した平面姿勢推定手法と多眼カメラへの拡張. 情報処理学会研究報告 (コンピュータビジョンとイメージメディア 2005-CVIM-151), No. 112, pp. 131–138, 2005.
- 17) 内田秀雄, 杉本茂樹, 奥富正敏. ステレオ時系列画像を用いた直接法による高速・高精度モーション推定. 電子情報通信学会論文誌, Vol. J92-D, No.8, pp. 1414–1424, 2009.